

König, Johannes; Lebens, Morena

**Classroom Management Expertise (CME) von Lehrkräften messen:  
Überlegungen zur Testung mithilfe von Videovignetten und erste  
empirische Befunde**

*Lehrerbildung auf dem Prüfstand 5 (2012) 1, S. 3-28*



Quellenangabe/ Reference:

König, Johannes; Lebens, Morena: Classroom Management Expertise (CME) von Lehrkräften messen: Überlegungen zur Testung mithilfe von Videovignetten und erste empirische Befunde - In: Lehrerbildung auf dem Prüfstand 5 (2012) 1, S. 3-28 - URN: urn:nbn:de:0111-pedocs-147296 - DOI: 10.25656/01:14729

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-147296>

<https://doi.org/10.25656/01:14729>

in Kooperation mit / in cooperation with:



[www.vep-landau.de](http://www.vep-landau.de)

#### Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

#### Kontakt / Contact:

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Mitglied der

  
Leibniz-Gemeinschaft



**Herausgeber**

Rainer Bodensohn, Andreas Frey und Reinhold S. Jäger  
Bürgerstraße 23, 76829 Landau/Pfalz  
Telefon: +49 6341 280 32165, Telefax: +49 6341 280 32166

**Verlag**

Empirische Pädagogik e. V.  
Bürgerstraße 23, 76829 Landau/Pfalz  
Telefon: +49 6341 280 32180, Telefax: +49 6341 280 32166  
E-Mail: [info@vep-landau.de](mailto:info@vep-landau.de)  
Homepage: <http://www.vep-landau.de>

**Umschlaggestaltung**

© Harald Baron

**Druck**

DIFO Bamberg

*Alle Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung, werden vorbehalten. Kein Teil des Werks darf in irgendeiner Form (durch Fotografie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung des Verlags reproduziert oder unter Verwendung elektronischer Systeme verbreitet werden.*

ISSN 1867-2779

© Verlag Empirische Pädagogik, Landau 2012

---

---

## Inhalt

---

---

<b>Editorial</b> .....	1
<b>Originalarbeiten</b>	
König, J. & Lebens, M.: <i>Classroom Management Expertise</i> (CME) von Lehrkräften messen: Überlegungen zur Testung mithilfe von Videovignetten und erste empirische Befunde .....	3
Oesterhelt, V., Gröschner, A., Seidel, T. & Sygusch, R.: Pädagogische Vorerfahrungen und Kompetenzeinschätzungen im Kontext eines Praxissemesters – Domänenspezifische Betrachtungen am Beispiel der Sportlehrerbildung .....	29
Schneider, C., Bodensohn, R. & Müller, A.: Das Zusammenspiel pädagogischer Handlungskompetenz und fachdidaktischer Kompetenz in Selbsteinschätzungen angehender Lehrer mathematisch-naturwissenschaftlicher Fächer in der ersten Phase der Ausbildung .....	47
Weiß, S., Schramm, S. & Kiel, E.: Was sollen Grundschullehrkräfte können? Die Sicht von Lehrkräften und Ausbildungspersonen .....	81
<b>Vorstellungen</b>	
Terhart, E., Schulze-Stocker, F., Kunina-Habenicht, O., Dicke, T., Förster, D., Lohse-Bossenz, H., Gößling, J., Kunter, M., Baumert, J. & Leutner, D.: Bildungswissenschaftliches Wissen und der Erwerb professioneller Kompetenz in der Lehramtsausbildung. Eine Kurzdarstellung des BilWiss-Projekts .....	96
<b>Impressum</b> .....	107

---

---

## Contents

---

---

### Articles

- König, J. & Lebens, M.:  
Measuring teachers' *Classroom Management Expertise* (CME):  
On the testing via video-vignettes and first empirical findings ..... 3
- Oesterhelt, V., Gröschner, A., Seidel, T. & Sygusch, R.:  
Pedagogical pre-experiences and perceived competences in the context  
of a practical internship – domain-specific considerations in physical  
education teacher education ..... 29
- Schneider, C., Bodensohn, R. & Müller, A.:  
The interplay of teacher education, students' self-ratings of general  
pedagogical competence and pedagogical-content-related competence ... 47
- Weiß, S., Schramm, S. & Kiel, E.:  
What should elementary school teachers know? The sight of teachers  
and teachers' trainers ..... 81

---

## Originalarbeiten

---

Johannes König und Morena Lebens

### *Classroom Management Expertise* (CME) von Lehrkräften messen: Überlegungen zur Testung mithilfe von Videovignetten und erste empirische Befunde

**Zusammenfassung:** Klassenführung (*Classroom Management*) gilt übereinstimmend als zentrale Basisdimension der Unterrichtsqualität und stellt an Lehrkräfte hohe berufliche Anforderungen. Lehrkräfte benötigen für die Bewältigung solcher Anforderungen *Expertise* (pädagogisches Wissen und Können). Im vorliegenden Beitrag werden Überlegungen zur Messung der *Classroom Management Expertise* (CME) von Lehrkräften mit Videovignetten dargestellt. Dazu wird die wissensgesteuerte Verarbeitung von Unterricht anhand von drei Dimensionen kognitiver Anforderungen konkretisiert: *Genauigkeit der Wahrnehmung*, *Holistische Wahrnehmung* und *Rechtfertigung einer Handlung*. Am Beispiel einer Videovignette und prototypischer Testaufgaben, die diese drei kognitiven Anforderungsdimensionen exemplarisch abbilden, sowie unter Verwendung einer Stichprobe ( $n = 108$ ) von Lehramtsstudierenden, Referendaren und berufstätigen Lehrpersonen, die Experten-Novizen-Vergleiche ermöglicht, erfolgt eine empirische Untersuchung von CME. Die Ergebnisse verdeutlichen erwartungskonform, dass berufstätige Lehrkräfte besser als Referendare und diese wiederum besser als Lehramtsstudierende in der Lage sind, die in der Videovignette gezeigte Unterrichtssituation genauer und holistischer wahrzunehmen. Tendenziell sind die Referendare eher in der Lage, eine beobachtete Handlung zu rechtfertigen. Die Ergebnisse werden unter Bezugnahme auf Erkenntnisse der Expertiseforschung diskutiert.

**Schlagwörter:** Klassenführung – Lehrerexpertise – Lehrerkompetenzen – Leistungstest – Videovignette

### Measuring teachers' *Classroom Management Expertise* (CME): On the testing via video-vignettes and first empirical findings

**Abstract:** Classroom management is a central dimension of teaching quality and a professional task teachers have to master. For this they need expertise (pedagogical knowledge and performance). In this article we outline how *Classroom Management Expertise* (CME) of teachers can be measured by using video-vignettes. To do this, we conceptualize the knowledge-based reasoning of classroom situations across three cognitive dimensions: *accuracy of noticing*, *holistic noticing*, and *justification of action*. These dimensions are operationalized into prototypical test-items of a video-vignette. A sample ( $n = 108$ ) of pre-service teachers (first and second phase of teacher education) and in-service teachers is used to allow expert-novice-comparisons and to empirically investigate CME. According to our assumptions, findings show in-service teachers outperform pre-service teachers of the second phase, and they in turn outperform pre-service teachers of the first phase when looking at how accurate and how holistic they notice the presented classroom situation. However, pre-service teachers of the second phase show a tendency to outperform the other two groups when asked to justify an action. Findings are discussed and related to research on teacher expertise.

**Key words:** achievement test – classroom management – teacher competencies – teacher expertise – video-vignette

## 1. Einleitung

Verschiedene empirische Untersuchungen haben in den vergangenen Jahren zu einem erheblichen Fortschritt in der standardisierten Erfassung von kognitiven Elementen professioneller Kompetenz angehender wie berufstätiger Lehrkräfte<sup>1</sup> beigetragen und Erkenntnisse zur Wirksamkeit der Lehrerbildung geliefert (v.a. Blömeke, Kaiser & Lehmann, 2010a, b; Blömeke et al., 2011; König & Seifert, 2012; Kunter et al., 2011). Diese Untersuchungen besitzen die Gemeinsamkeit, sich auf ein allgemeines Modell professioneller Kompetenz von Lehrkräften zu beziehen (Baumert & Kunter, 2006, 2011; Blömeke, Kaiser & Lehmann, 2010a, b): Auf der einen Seite finden sich Bereiche des professionellen Wissens von angehenden bzw. berufstätigen Lehrkräften; ihnen gegenüber stehen motivational-affektive Charakteristika. Das Professionswissen wird in Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen gegliedert. Als Facetten werden zumindest deklaratives Wissen („Wissen, dass ...“) und prozedurales Wissen („Wissen, wie ...“) unterschieden. Die in den Modellen professioneller Kompetenz enthaltenen Facetten werden im Sinne des Kompetenzbegriffs (Weinert, 2001) als Voraussetzungen der Lehrkräfte verstanden, um zentrale berufliche Anforderungen erfolgreich bewältigen zu können.

Zur Erfassung des Professionswissens werden meist Papier-Bleistift-Verfahren angewendet, um eine forschungsökonomische Testung an großen Stichproben zu realisieren. Vermutlich bewährt sich dieser forschungsmethodische Zugang, solange vor allem das theoretisch-formale bzw. konzeptuell-analytische Wissen von Lehrkräften Gegenstand der Testung ist. Eine besondere Herausforderung hingegen besteht darin, Wissen zu erfassen, das besonders eng mit der tatsächlichen Bewältigung spezifischer beruflicher Anforderungen verknüpft ist und somit nahe an den in der empirischen Bildungsforschung genutzten Kompetenzbegriff heranreicht, mit dem Kompetenzen definiert werden als „*kontextspezifische kognitive Leistungsdispositionen*, die sich funktional auf Situationen und Anforderungen in bestimmten *Domänen* beziehen“ (Klieme & Leutner, 2006, S. 879).

Folgt man zum Beispiel dem Erkenntnisstand der Forschung zur Lehrerexpertise, so wäre zu fordern, dass die Testung von Professionswissen noch stärker als es verschiedene derzeit verfügbare Testverfahren realisieren (können), kontext- und situationsbezogen erfolgt (vgl. für das pädagogische Wissen z. B. König, 2010). Damit wäre verschiedenen, zum Teil auch grundsätzlichen Einwänden an standardisierten Messungen von Lehrerkompetenzen Rechnung getragen, etwa der Kritik, ob mit Papier-Bleistift-Verfahren nicht einfach nur träges Wissen ge-

---

<sup>1</sup> Um den Lesefluss zu erleichtern, wird nachfolgend für Lehrerinnen und Lehrer, Referendarinnen und Referendare usw. nur die männliche Form verwendet. Es sind jedoch stets beide Geschlechter gemeint, wenn nicht anders gekennzeichnet.

messen werde, das zwar in der universitären Ausbildung vermittelt und erworben wird, bei der Ausübung des Berufs jedoch bedeutungslos ist (z. B. Neuweg, 2011). Auch könnte es noch besser gelingen, jene Kompetenzen direkt zu messen, die von angehenden Lehrkräften in einem fortgeschrittenen Stadium (z. B. in der zweiten Ausbildungsphase) auf der Basis ihrer ersten praktischen Handlungserfahrungen erworben werden (z. B. Identifikation von bedeutungsvollen situativen und kontextualisierten Elementen) und anhand derer sie sich von angehenden Lehrkräften, die sich noch am Anfang ihrer Ausbildung befinden, unterscheiden (vgl. König, 2010).

Der vorliegende Beitrag setzt an diesem Problem an. Am Beispiel des pädagogischen Wissens und Könnens von Lehrkräften mit Blick auf die zentrale Anforderung der Klassenführung werden Überlegungen dargestellt, die von einer bisher realisierten Papier-Bleistift-Testung hin zur Notwendigkeit einer videobasierten Testung führen. Diese Überlegungen werden mit Beispiel-Testaufgaben, die an eine Videovignette gekoppelt sind, konkretisiert und empirisch über Antworten von Experten- und Novizen-Lehrern vergleichend untersucht.

### 1.1 Pädagogisches Wissen und Können als Basis von Klassenführungsexpertise

Neben rein fachlichen Anforderungen im Lehrerberuf (z. B. Strukturierung der Inhalte eines Faches) stehen Lehrkräfte vor zentralen Anforderungen, die, auch wenn sie z. T. fachspezifisch ausgestaltet werden müssen, für alle Fächer relativ gleichförmig sind. Eine dieser Anforderungen ist die Klassenführung (*Classroom Management*). Das Führen einer Schulklasse stellt hohe berufliche Anforderungen an Lehrkräfte, für deren erfolgreiche und professionelle Bewältigung Expertise notwendig ist (Doyle, 1985, 2006; Evertson & Weinstein, 2006; Helmke, 2009). Aus diesem Grund wird für ein erfolgreiches *Classroom Management* ein fächerübergreifendes, pädagogisches Wissen und Können vorausgesetzt (Bromme, 1992, 1997; Doyle, 1985; Fennema & Franke, 1992; Grossmann, 1990; König, Blömeke, Paine, Schmidt & Hsieh, 2011; Shulman, 1987).

Schon lange wird die kognitive Fundierung der Fähigkeit von Lehrkräften, eine Klasse zu führen, betont (Doyle, 1985, 2006; Evertson & Weinstein, 2006): Erfolgreiche Klassenführung hängt grundsätzlich mit der Fähigkeit von Lehrkräften zusammen, die relevanten Ereignisse des komplexen unterrichtlichen Geschehens zu erkennen und zu interpretieren. Diese intellektuelle Fähigkeit der Lehrkräfte fußt auf ihrem Wissen bzw. den kognitiven Schemata zur Klassenführung sowie ihren kognitiven Repräsentationen situativer Handlungen im unterrichtlichen Kontext. Dabei geht es nicht um eine unstrukturierte Ansammlung von ‚Tricks‘ oder spezifischen Verhaltensreaktionen (z. B. „*don't smile before christmas*“), sondern um einen „*intellectual framework*“ des Managements (Doyle, 1985,



S. 33), den Lehrkräfte zu erwerben haben. Hierzu zählt unter anderem propositionales Wissen zum Unterricht bzw. unterrichtlichen Kontext und Verfahren der angemessenen Klassenführung (vgl. auch Copeland, 1987).

Wissen über Klassenführung als Merkmal professioneller Kompetenz von Lehrkräften ist im deutschsprachigen Raum bislang zum Gegenstand einer auf standardisierten Testverfahren beruhenden Kompetenzmessung gemacht worden. Die Verfahren seien im Folgenden kurz beschrieben, bevor auf den theoretischen Rahmen der vorliegenden Untersuchung eingegangen wird.

- (1) Im Rahmen der *Teacher Education and Development Study – Learning to Teach Mathematics* (TEDS-M; Blömeke et al., 2010a, b) wurde ein Papier-Bleistift-Test zur Erfassung von pädagogischem Wissen von angehenden Lehrkräften im letzten Jahr ihrer Ausbildung entwickelt, der unter anderem Wissen zur Klassenführung erfasst (König & Blömeke, 2009). Inhaltlich berücksichtigt wurde dabei Wissen zu störungspräventiver Unterrichtsführung (z. B. Planungsaspekte, konkretes Lehrerverhalten) und zur effektiven Nutzung der Unterrichtszeit (z. B. Unterrichtsroutinen). Der Test wurde für den internationalen Vergleich der Lehrerbildung in den USA, Taiwan und Deutschland konzipiert und an repräsentativen Länderstichproben angehender Mathematiklehrkräfte im letzten Jahr ihrer Ausbildung eingesetzt (König et al., 2011).
- (2) Ferner wurde im Forschungsprogramm *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz* (COACTIV; Kunter et al., 2011) ein Papier-Bleistift-Test zur Erfassung pädagogisch-psychologischen Wissens entwickelt, der Wissen von angehenden Lehrkräften in der zweiten Ausbildungsphase (Referendariat bzw. Vorbereitungsdienst) erfasst, darunter auch Wissen zur Klassenführung (Voss & Kunter, 2011). Speziell für die Erfassung von Wissen zur Klassenführung wurden Videoclips verwendet, die als Stamm der Testaufgaben fungierten.
- (3) Videogestützte Verfahren zur Erfassung von Unterrichtsbeobachtungskompetenz von angehenden Lehrkräften (Projekt *Observe*, vgl. Seidel, Blomberg & Stürmer, 2010; Projekt *Kompetenzen zur Analyse der Lernwirksamkeit von naturwissenschaftlichem Grundschulunterricht – Theoretische Modellierung und empirische Erfassung*, vgl. Gold et al., 2011) schließen auch die Beurteilung von Aspekten der Klassenführung mit ein. Diese Verfahren werden bislang insbesondere an Stichproben von Lehramtsstudierenden angewendet bzw. entwickelt. Das Projekt *Observe* zielt darauf, professionelle Wahrnehmung als Bestandteil von Lehrerexpertise zu messen, wobei diese Wahrneh-

mung in Anlehnung an van Es und Sherin (2008) in die beiden Komponenten „Wahrnehmung von Unterrichtskomponenten“ (*noticing*) und „wissensgesteuerte Verarbeitung von Unterricht“ (*knowledge-based reasoning*) unterteilt wird (Seidel et al., 2010, S. 297). Seidel et al. (2010) beziehen die *noticing*-Komponente mit ihrem Instrument „Observer“ auf *allgemeine* Merkmale qualitativollen Unterricht (Zielorientierung, Lernbegleitung, Lernatmosphäre). Mit der Komponente des *knowledge-based reasoning* werden ferner drei Ebenen wissensgesteuerter Verarbeitung bei der professionellen Beobachtung unterschieden: Unterrichtssituationen und ihre Wirksamkeit unter Einbezug theoretischen Wissens beschreiben, erklären, vorhersagen.

- (4) In der DFG-Studie *Längsschnittliche Entwicklung der Kompetenzen von Junglehrkräften: Follow-Up zur internationalen Vergleichsstudie TEDS-M* (TEDS-FU; Blömeke, Kaiser & König, 2009) werden derzeit, aufbauend auf die TEDS-M-Tests zur Erfassung von mathematischem Fachwissen, mathematikdidaktischem Wissen und pädagogischem Wissen, *fachspezifische* Videovignetten entwickelt und an Junglehrkräften mit dem Unterrichtsfach Mathematik validiert.

## 1.2 Theoretischer Rahmen der Untersuchung

Der dargelegte Forschungsstand macht deutlich, dass eine spezifische Modellierung von Klassenführungscompetenz bzw. Messung von *Classroom Management Expertise* von Lehrkräften als Maß des pädagogischen Wissens und Könnens von angehenden bzw. berufstätigen Lehrkräften vorwiegend auf der Ebene von Papier-Bleistift-Verfahren angegangen wird. Während das TEDS-M-Testinstrument wie auch das im Rahmen von COACTIV-R verwendete Instrument Wissen zur Klassenführung als Subskala explizit theoretisch ausweisen und empirisch abbilden können, ist eine solche Spezifikation im Projekt *Observe* nicht erkennbar. Nötig ist, die bislang begonnene spezifische Erfassung von Wissen zur Klassenführung messmethodisch so zu erweitern, dass eine Testung jener kognitiven Anforderungen, die an Lehrkräfte im Bereich der Klassenführung gestellt werden, nicht nur auf einer deklarativen Ebene, sondern in stärkerem Maße kontext- und situationsbezogen erfolgt.

In der nachfolgenden Untersuchung wird, auch in Anlehnung an die Überlegungen aus COCATIV-R (vgl. Voss & Kunter, 2011, S. 202), davon ausgegangen, dass die Darbietung von Videovignetten einen angemessenen Stimulus im Testaufgabenstamm bieten kann, der der Forderung einer stärkeren kontext- und situationsbezogenen Testung besser gerecht wird als es allein mit der Verwendung von in Textform dargebotenen Unterrichtssituationen der Fall wäre. Darüber

hinaus gilt es jedoch, den dadurch entstehenden Mehrwert solcher Testaufgabenstämme zu spezifizieren.

Aktuell wird der Einsatz von Videovignetten bei Testungen zur Erfassung von Lehrerkompetenzen favorisiert, weil mit dem methodischen Zugang allgemein Nähe zum tatsächlichen Handeln im Lehrerberuf gegeben sei und die Komplexität des unterrichtlichen Geschehens abgebildet werde (z. B. Neuweg, 2011; Schaper, 2009; Voss & Kunter, 2011). Diese allgemeinen Vorteile sehen wir auch; zudem nehmen wir jedoch an, dass es einer weiteren Präzisierung bedarf, worin der spezifische Vorteil einer Erweiterung von Papier-Bleistift-Testungen durch die Verwendung von Videovignetten besteht, um pädagogisches Wissen und Können von angehenden bzw. berufstätigen Lehrkräften zu messen. Eine Messung von Expertise zur Klassenführung mithilfe eines videobasierten Verfahrens, so die hier vertretene Position, sollte nicht nur in einen methodischen, sondern auch in einen inhaltlichen Begründungszusammenhang gestellt werden. Um dies zu klären, sind kognitive Anforderungen zu bestimmen, die an die Probanden während der Testung gestellt und über bestimmte Testaufgaben operationalisiert werden. In Anlehnung an den unter Punkt 1.1 dargelegten Forschungsstand, Erkenntnisse der Forschung zur Lehrerexpertise und Klassifikationen zur Repräsentation von Lehrwissen (u. a. Baumert & Kunter, 2006, 2011; Bromme, 1992, 1997, 2001; Fenstermacher, 1994; König, 2010; Neuweg, 2008, 2011; Plöger & Scholl, 2011; Putnam & Borko, 2000) beschreiben und begründen wir deshalb im Folgenden eine Konkretisierung der wissensgesteuerten Verarbeitung von Unterricht anhand von drei Dimensionen kognitiver Anforderungen:

- Genauigkeit der Wahrnehmung,
- Holistische Wahrnehmung: Rekonstruktion und Antizipation einer Handlung und
- Rechtfertigung einer Handlung.

### 1.2.1 Genauigkeit der Wahrnehmung

Die Expertiseforschung belegt, dass Experten aufgrund ihrer „kategorialen Wahrnehmung“<sup>2</sup> relevante Details wesentlich besser und genauer erinnern sowie wiedergeben können als Novizen. Experten zeichnen sich zudem durch eine holistische Informationsentnahme aus, welche zu der Detailwahrnehmung hinzutritt. Das ermöglicht ihnen einen ganzheitlicheren Zugang zum unterrichtlichen Ge-

---

<sup>2</sup> „Damit sind die Schnitte gemeint, mit denen die gegebenen Sachverhalte oder ein Ereignisstrom kognitiv in Einheiten unterteilt werden. Durch die verfügbaren Begriffe (das können auch nichtsprachlich gefasste Begriffe sein) wird eine gegebene Situation für den Handelnden strukturiert“ (Bromme, 1992, S. 42).

schehen wie auch eine komplexere Wahrnehmung der Situation (vgl. Carter et al., 1987; Plöger & Scholl, 2011). Für Novizen sind dagegen zunächst einmal viele Einzelinformationen wichtig (z. B. Sabers, Cushing & Berliner, 1991), da sie sich in der Komplexität der unterrichtlichen Situation verlieren. Oder aber sie erinnern Details, die funktional für die Klassenführung nicht oder nur teilweise relevant sind.

### 1.2.2 Holistische Wahrnehmung: Rekonstruktion und Antizipation einer Handlung

Aus der Expertiseforschung ist ferner bekannt, dass Experten bei der Wahrnehmung von Unterrichtssituationen sogleich angeregt werden, an Handlungsoptionen bzw. die Steuerung des weiteren Unterrichtsablaufs zu denken (z. B. Gruber, Harteis & Rehl, 2006; Shavelson & Stern, 1981). Dies hängt damit zusammen, dass ihnen zahlreiche typische, also ähnliche Unterrichtsszenarien wohl vertraut sind, in die sie eine betrachtete Unterrichtssituation einordnen können. Ihr durch eine holistische Wahrnehmung (Bromme, 1992, 2001) geprägter Blick auf eine dargebotene Unterrichtssituation verbindet sich unmittelbar mit einem Problemlöseverhalten, das Klarheit über die zeitliche und räumliche Orientierung herstellt und Antworten gibt auf Fragen, wie es z. B. zu dieser Situation gekommen ist, welche Zielstellung dem betrachteten Unterrichtsausschnitt unterliegt oder in welche Richtung der Unterricht daher verlaufen wird. Novizen hingegen, die bislang nur wenig eigenen Unterricht geplant und gehalten haben, und die bei der Durchführung von eigenem Unterricht rezeptartig und unflexibel an ihrer vorab ausführlich erstellten Planung festhalten (vgl. Leinhardt & Greeno, 1986), dürfte es schwerfallen, aus der betrachteten Situation heraus die weitere, typische Handlungsstruktur des Unterrichts flüssig vorherzusehen bzw. die vorhergehende Handlungsstruktur zu rekonstruieren und an alternative Handlungen im Unterricht zu denken sowie diese zu vergleichen und zu bewerten. Grund hierfür ist die bei ihnen noch nicht entwickelte kognitive Struktur und das noch nicht oder kaum vorhandene Fallwissen, welches bei Experten reichhaltig und erfahrungsgesättigt vorliegt. Deren Wahrnehmung ist in Form von zentralen Konzepten und Ereignisschemata organisiert, welche zu großen Teilen auf Erfahrung beruhen. An diese Konzepte und Schemata sind dann „Einzelfälle, episodische Einheiten oder Sequenzen von Episoden (Skripts) angedockt“ (Baumert & Kunter, 2006, S. 483). Angesichts dieser Unterschiede sollten Expertenlehrpersonen besser in der Lage sein, den weiteren Verlauf eines betrachteten Videoclips vorherzusehen, ihn in das gesamthafte zeitliche und räumliche Unterrichtsgeschehen einzuordnen, sowie dafür von den im Videoclip dargelegten Informationen insgesamt reichhaltiger Gebrauch machen als Novizen.

### 1.2.3 Rechtfertigung einer Handlung

Warum ist ... geschehen? Warum positioniert sich die Lehrerin vor der Tafel, warum diese Sitzordnung? Zur Beantwortung solcher Fragen im Zusammenhang mit der funktionalen Interpretation von Unterrichtssituationen ist handlungsrechtfertigendes bzw. handlungssteuerndes Wissen (Neuweg, 2008, S. 17) von Bedeutung, das neben den zuvor genannten Anforderungsdimensionen eine weitere Dimension darstellt. Im Schulalltag wird dieses Wissen vermutlich eher selten expliziert, es ist jedoch im Zuge der Expertisierung verdichtet und grundsätzlich abrufbar (Bromme, 1992). Neben einer quantitativen Wissensverdichtung zeichnet sich die Rechtfertigung einer Handlung im Vergleich zur holistischen Wahrnehmung durch eine qualitative Entwicklung von der Repräsentation hin zur Transformation des Wissens aus. Erst die Fähigkeit zur Wissenstransformation ermöglicht es Lehrkräften, situationsspezifische Schlussfolgerungen zu ziehen und diese verbal zu explizieren. Durch die Betrachtung von Videovignetten sollten Expertenlehrer stärker als Novizen angeregt werden, relevante Schlussfolgerungen zu ziehen, die im Zusammenhang mit solchem Wissen stehen. Die Differenzierung in Sicht- und Tiefenstruktur von Unterricht (Oser & Patry, 1994; Klieme, 2006) kann hier eine wichtige Rolle spielen, z. B. in dem Sinne, dass sich Novizen an der Sichtstruktur des Unterrichts orientieren („die Schüler/innen arbeiten in Gruppen“), sich also auf das direkt Beobachtbare beziehen, während Experten darüber hinaus das Sichtbare unmittelbar funktional auf weitere Elemente des Unterrichts, z. B. seine Zielsetzungen, beziehen („Gruppenarbeit statt Stillarbeit, weil die Aufgabenstellung arbeitsteilig und leistungsdifferenziert erfolgen soll“), und damit das nur interpretativ Zugängliche mit einbeziehen. Zusammenfassend kann die holistische Wahrnehmung als perzeptiver, die Rechtfertigung einer Handlung als interpretativer Vorgang bezeichnet werden.

## 2. Fragestellung

Kaum eine andere Anforderung im Lehrerberuf ist neben einer kognitiven Fundierung so stark an Erfahrung gebunden wie das Führen einer Klasse. Aus diesem Grund nehmen wir an, dass eine Modellierung von *Classroom Management Expertise* über den Vergleich von Experten und Novizen gelingt. Wir stellen somit die folgenden Fragen:

- (1) Unterscheiden sich mehrjährig berufstätige Lehrkräfte (Experten) von angehenden Lehrkräften, d. h. von Referendaren bzw. Lehramtsstudierenden, in ihrer Wahrnehmung von Unterrichtssituationen, in denen Expertise (d. h. Wissen und Können) zur Klassenführung von Bedeutung ist?

- (2) Zeigen sich dabei Unterschiede in den Dimensionen *Genauigkeit der Wahrnehmung*, *Holistische Wahrnehmung* und *Rechtfertigung einer Handlung*, die zugunsten erfahrener Lehrkräfte ausfallen?

Während die Anforderung der ersten Dimension vor allem auf das Betrachtete beschränkt bleibt, kann bei den beiden anderen Anforderungen darüber hinaus auch das Abrufen externen Wissens eine Rolle spielen (z. B. didaktische Funktionen von Gruppenarbeit kennen). Angehenden Lehrkräften wird dies, so ist zu vermuten, dann gelingen, wenn z. B. theoretisches, ausbildungsnahes Wissen vorhanden ist, das sie mit dem Betrachteten verknüpfen können (z. B. „Gruppenarbeit kann kooperatives Lernen fördern“). Expertenlehrer werden hierbei vor allem den Bezug zur Stunde als Ganzes als Richtgröße haben und das Betrachtete stärker situativ verorten. Insofern sollten sich nicht nur „bessere“ und „schlechtere“ Antworten, sondern zudem unterschiedliche Qualitäten der Beantwortung zeigen.

### 3. Methode

#### 3.1 Stichprobe

Zur Bearbeitung der Fragestellung wurden drei unterschiedliche Gruppen von angehenden bzw. berufstätigen Lehrkräften einbezogen, um kontrastive Vergleiche im Sinne des Experten-Novizen-Paradigmas hinsichtlich ihrer kognitiven Bearbeitung von videobasierten Testaufgaben zur Klassenführung zu ermöglichen: Unter angehenden Lehrkräften wurde einerseits eine Gruppe von Lehramtsstudierenden der Universität zu Köln, die sich im Hauptstudium befanden, also über erste Praxiserfahrungen in Schulen verfügten ( $n_{\text{Stud}} = 19$ ), sowie andererseits eine Gruppe von Referendarinnen und Referendaren, die bereits regelmäßig Unterricht erteilt hatten ( $n_{\text{Ref}} = 73$ ), einbezogen. Ferner beteiligten sich berufstätige Lehrkräfte mit mindestens zwei, im Durchschnitt jedoch 17 Jahren Berufserfahrung ( $M = 17.1$ ;  $SD = 12.4$ ;  $\text{Min.} = 2$ ,  $\text{Max.} = 40$ ) an der Testung ( $n_{\text{Lehr}} = 16$ ). Alle Probandinnen und Probanden wurden in Köln bzw. in der Kölner Region rekrutiert und unter einheitlicher Instruktion befragt bzw. getestet.<sup>3</sup> Tabelle 1 enthält Kennwerte über die Zusammensetzung der Stichproben. Da es sich um eine Pilotstudie handelt, die aus forschungsökonomischen Gründen mit vergleichsweise wenig Aufwand durchgeführt werden musste, wurden Gelegenheitsstichproben verwendet. In Bezug auf die Lehramtsstudierendenstichprobe wie auch die Stichprobe von Referendaren war es jedoch möglich, die Testung in einer universitären Lehrveranstaltung bzw. in einem Ausbildungsseminar der zweiten Phase durch-

---

<sup>3</sup> Allen Teilnehmerinnen und Teilnehmern der Studie sowie den Personen, die die Datenerhebung unterstützt haben, sei an dieser Stelle herzlich gedankt.

zuführen. Die Gefahr einer individuellen Selbstselektion kann somit für diese beiden Befragungsgruppen weitgehend ausgeschlossen werden.

Tabelle 1: Zusammensetzung der Stichprobe, gegliedert nach Expertise-Gruppe

	n	Anteil weiblich (in %)	Alter M (SD)	Semester M (SD)
Lehramtsstudierende im Hauptstudium	19	61.8	24.0 (1.0) <sup>a</sup>	6.3 (2.3)
Referendarinnen bzw. Referendare	73	63.2	30.2 (6.4)	-
Berufstätige Lehrpersonen	16	53.8	46.5 (9.3)	-

<sup>a</sup> Schätzung.

### 3.2 Erhebungsinstrument

Die verwendete Videovignette besteht aus einem kurzen Unterrichtsvideoausschnitt als Testaufgabenstamm und vier sich daran anschließenden Testfragen. Die Probanden durften den Videoausschnitt nur einmal betrachten und erst im Anschluss die Testfragen lesen. Der Unterrichtsvideoausschnitt ist der Internetseite „Guter Unterricht“ des Landesinstituts für Lehrerbildung und Schulentwicklung Hamburg entnommen, das verschiedene Unterrichtsvideos auf der Webseite kostenlos anbietet ([www.guterunterricht.de](http://www.guterunterricht.de)). Bei dem gewählten Ausschnitt handelt es sich um den vierten Videoclip eines Videos einer Ethikunterrichtsstunde in einer 10. Klasse mit dem Stundenthema „Suizid“. Der Videoclip trägt den Titel „4. Phasenübergang: Selbstständige zu gemeinsamer Arbeit“ und dauert 81 Sekunden. Nur dieser Videoclip wurde den Probanden gezeigt, wobei der zu Beginn erscheinende Titel vom Clip entfernt und der so bearbeitete Clip separat und nicht von der Webseite aus den Probanden vorgeführt wurde.<sup>4</sup>

Im Videoausschnitt ist eine an Gruppentischen arbeitende Schulklasse zu sehen. An jedem Gruppentisch arbeiten ca. vier Schüler mit einem Laptop pro Gruppe. Es ertönt ein akustisches Signal (Glocke), woraufhin die Lehrerin ins Blickfeld der Kamera gerät. Die Lehrerin instruiert die Schüler, die Gruppenarbeitsphase zu schließen, eine Entspannungsübung durchzuführen und kündigt schließlich den Beginn der Präsentationsphase an. Der Videoausschnitt macht somit den Phasenwechsel (*transition*) zum Gegenstand, eine zentrale Dimension der Klassenführung, der bei einem effektiven *Classroom Management* zügig und reibungslos verlaufen soll (Kounin, 1976).

<sup>4</sup> Der Videoclip ist verfügbar unter: <http://www.guterunterricht.de/guterunterricht.de/Video1.html>.

Die vier Testfragen (drei offene, eine geschlossene; vgl. Abbildung 1) decken die unter Punkt 1.2 dargelegten kognitiven Anforderungen ab. Die erste Frage zielt auf Aspekte der *Genauigkeit der Wahrnehmung* (Dimension 1) sowie auf Aspekte der *Rechtfertigung einer Handlung* (Dimension 3). In Hinblick auf die Genauigkeit der Wahrnehmung geht es um die Frage, ob die Probanden lediglich die Sichtstruktur des Unterrichts wahrnehmen in dem Sinne, dass sie erkennen, dass Gruppenarbeit durchgeführt wird und dafür die Jugendlichen an Gruppentischen sitzen; oder ob sie darüber hinaus auch erkennen, dass die Sitzordnung eine Funktion für die sichtbaren Tätigkeiten der Jugendlichen hat. Diese sind operationalisiert über die drei Aspekte Kommunikation, Zugang zu einem Computer und Vorbereitung einer Präsentation in der Gruppe. Es geht hier also um die Wiedergabe bzw. Beschreibung von konkret Wahrgenommenem, das die Gruppenarbeit als für den Unterricht gewählte Form begründen hilft. Ferner enthält die Frage auch die kognitive Anforderung, die Gruppenarbeit mit einer allgemeineren Funktion oder Zielsetzung zu rechtfertigen, die so nicht direkt beobachtet werden kann (z. B. kooperatives Lernen, Verfolgen von sozialen Lernzielen). Dafür ist es erforderlich, eine Verknüpfung mit externem theoretischem Wissen (z. B. Lehrbuchwissen) zu Funktionen von Gruppenarbeit vorzunehmen. Schließlich werden auch Antworten dahingehend kodiert, ob sie auf eine Funktion Bezug nehmen, die über die konkrete Organisationsform der Gruppenarbeit hinausgeht: zum Beispiel, dass die Sitzordnung neben Gruppenarbeit auch Plenumsarbeit ermöglicht. Da die Bezugnahme auf eine solche Funktion einen besonders hohen Verknüpfungsgrad darstellt und auf eine globale Betrachtung der unterrichtlichen Situation zielt, zählen wir solche Antworten zur zweiten Dimension kognitiver Anforderungen (*Holistische Wahrnehmung*).

Die zweite Frage bildet ebenfalls die kognitive Dimension der *Holistischen Wahrnehmung* ab, jedoch über das Format einer geschlossenen Testaufgabe. Dabei geht es um die allgemeine Einordnung des Betrachteten in den zeitlichen Verlauf einer Unterrichtsstunde. Die korrekte Antwortoption C erfordert es, unterschiedliche Informationen aus dem betrachteten Video zusammenzuführen und die Wahrnehmung des komplexen Geschehens in Bezug auf den typischen zeitlichen Ablauf einer Unterrichtsstunde hin zu präzisieren.

Die Frage 3 erfasst wiederum Aspekte, die der ersten Dimension *Genauigkeit der Wahrnehmung* zuzuordnen sind. Die Lehrerin verwendet gleich mehrere Maßnahmen, um die Gruppenarbeitsphase zu beenden und die Präsentationsphase einzuleiten. Die Maßnahmen folgen sehr rasch aufeinander und sind von unterschiedlicher Qualität (Glocke, Aufforderung, den Computer zu schließen und die Stifte wegzulegen) und werden sozial unterschiedlich kommuniziert (Ansprache an die gesamte Lerngruppe, Ansprache an einen einzelnen Schüler, der beim Na-



men gerufen wird). Expertenlehrer dürften vertraut sein mit diesen Maßnahmen, sie also auch genauer wahrnehmen und anschließend besser erinnern können, während Novizen Schwierigkeiten haben dürften, die Vielfalt der Maßnahmen präzise wiedergeben zu können.

Die vierte Frage schließlich erfasst Aspekte, die wir verschiedenen Dimensionen zuordnen: Auf der Basis der betrachteten Unterrichtssituation geht es um die Antizipation, wie die Stunde weiter verlaufen wird. Einerseits ist auch hierfür wieder die Genauigkeit der Wahrnehmung relevant, weil z. B. erst bei Berücksichtigung von einzelnen Details vorheriger Instruktion der Lehrerin klar wird, was geplant ist (zwei Gruppen werden ihre Ergebnisse präsentieren, dabei kommt ein Evaluationsbogen zum Einsatz, den die anderen Gruppen auswerten werden). Andererseits ist auch dabei wieder externes Wissen bzw. Fallwissen gefragt, das der Expertenlehrperson helfen wird, den weiteren Verlauf funktional mit anderen wichtigen Aspekten zu verknüpfen (z. B. mit der didaktischen Funktion der Ergebnissicherung und der Gewährleistung eines gemeinsamen Lernstandes).

**Betrachten Sie das Video und versuchen Sie, die Fragen so gut wie möglich zu beantworten.**

1. Welche Funktion hat die Sitzordnung der Schülerinnen und Schüler für den gezeigten Unterricht?

2. Wann findet die gesehene Situation zeitlich betrachtet ungefähr statt?

*Kreuzen Sie bitte nur ein Kästchen an.*

- A. Am Anfang einer Unterrichtsstunde (d. h. während der ersten 5 Minuten). ☐
- B. Im ersten Drittel einer Unterrichtsstunde. ☐
- C. Im letzten Drittel einer Unterrichtsstunde. ☐
- D. Am Ende einer Unterrichtsstunde (d. h. während der letzten ca. 5 Minuten). ☐

3. Nennen Sie vier verschiedene Handlungsmaßnahmen (Stichworte), mit denen die Lehrerin gezielt die Aufmerksamkeit der Schüler auf sich richtet.

4. Was folgt unmittelbar nach dem betrachteten Ausschnitt?

Abbildung 1: Vier Testfragen zur Erfassung von *Classroom Management Expertise*

### 3.3 Datenauswertung

In die folgende Datenauswertung werden insgesamt 15 dichotome Items einbezogen, die wir aus den vier in Abbildung 1 dargestellten und eingesetzten Fragen beziehen. Ein Item beziehen wir aus Frage 2 (*Multiple-Choice*-Item; korrekte Antwort ist Option C). Aus Frage 1 beziehen wir vier und aus den Fragen 3 und 4 jeweils fünf Items. Diese resultieren aus einer Kodierung der auf diese offenen Fragen gegebenen Antworten entlang verschiedener Kriterien. Dabei wird wie folgt vorgegangen: Wenn das Kriterium (s. Spalte „Bezeichnung“) der gegebenen Antwort entnommen werden kann, wird der Code „1“ für „Ja, vorhanden“ erteilt; wenn das Kriterium der Antwort nicht entnommen werden kann, so wird der Code „0“ für „Nein, nicht vorhanden“ gegeben. Tabelle 2 sind Originalantworten zu entnehmen (s. Spalte „Antwortbeispiele“), die den jeweiligen Erwartungshorizont eines solchen Kriteriums illustrieren.

Tabelle 2: Variablen der Videovignette mit Antwortbeispielen und Dimensionszugehörigkeit

Frage	Variable	Bezeichnung	Antwortbeispiele	Dimension		
				1	2	3
1	V1_1	Gruppenarbeit	„Gruppenarbeit“ „Einteilung in Gruppen“ „Teamarbeit“	x		
	V1_2	Gruppenarbeit um zu kommunizieren, PC-Zugang zu haben und/oder um eine Präsentation vorzubereiten	„Dadurch wird die Kommunikation verbessert ...“ „Jeder hatte Zugang zu einem PC in seiner Gruppe“ „Die SuS sollen in Gruppen arbeiten, um Präsentationen vorzubereiten und sich dabei auszutauschen“	x		
	V1_3	Kooperatives Lernen bzw. andere Rechtfertigung	„Für die Sozialform der Gruppenarbeit. Ermöglicht kooperatives Lernen“ „Soziale Kompetenz fördern“			x
	V1_4	Verallgemeinerung der Funktion der Sitzordnung über die Gruppenarbeit hinaus	„Begünstigt offene Unterrichtsformen und Unterrichtsmethoden wie z. B. Gruppenarbeit“		x	
	V2	Zeitliche Verortung	Option C		x	
3	V3_1	Akustisches Signal	„Glocke“	x		
	V3_2	Stifte weglegen	„Stifte weglegen“	x		
	V3_3	Laptops schließen	„Laptops schließen“	x		
	V3_4	Einzelner Schüler wird angesprochen	„direkte Ansprache (Christian)“	x		

Frage	Variable	Bezeichnung	Antwortbeispiele	Dimension		
				1	2	3
4	V3_5	Raumposition der Lehrerin	„stellt sich so in den Raum, dass sie von allen wahrgenommen wird“	x		
	V4_1	Präsentation	„Die Präsentation der Schülerarbeiten“	x		
	V4_2	Schüler als Gruppe	„Gruppen präsentieren ihre Arbeitsergebnisse“	x		
	V4_3	Präzise Anzahl der präsentierenden Gruppen	„Die zwei Gruppen werden nacheinander ihre Arbeitsergebnisse vorstellen“	x		
	V4_4	Evaluation	„Ausfüllen eines Ergebnisbogens“	x		
	V4_5	Funktionale Verknüpfung	„Die Sicherung des Erarbeiteten“			x

Dimension 1 – Genauigkeit der Wahrnehmung, Dimension 2 – Holistische Wahrnehmung, Dimension 3 – Rechtfertigung einer Handlung.

Die entsprechenden, vorwiegend niedrig-inferenten Leitfäden zur Kodierung dieser Kriterien wurden einerseits theoriegeleitet (vgl. dazu unsere Ausführungen unter Punkt 1.2 und 3.2), andererseits datengeleitet unter Verwendung einzelner, zufällig gezogener Fragebögen der Probanden entwickelt. Um die Qualität der so entwickelten und fertig gestellten Leitfäden zu prüfen, wurden die Antworten auf die drei offenen Testfragen von rund einem Drittel der bis dahin noch nicht betrachteten Fragebögen (38 von 108 Bögen) von zwei unabhängig arbeitenden und geschulten Ratern doppelt kodiert. Als Maß der Interrater-Reliabilität wurde Cohen's Kappa berechnet (Wirtz & Caspar, 2002). Werte größer 0.75 gelten dabei als sehr gute Übereinstimmung. Die für unser Anliegen berechneten Kappa-Werte variieren zwischen 0.67 und 1.0 mit einem Mittelwert von  $M = 0.82$  ( $SD = 0.11$ ) für die 14 Indikatoren (Variablen V1\_1 bis V4\_5, ohne V2; s. Tab. 2). Unterhalb der Grenze von 0.75 liegen lediglich zwei Testitems, sodass die für die Kodierung entwickelten Leitfäden als bewährt gelten können. Bei fehlender Übereinstimmung wurden in gemeinsamer Diskussion – auch unter Hinzuziehung von Experten – Einigungen erzielt. Schwierig zu kodierende Antworten wurden als „Grenzfälle“ dokumentiert, um im Anschluss ein konsistentes Vorgehen mit ähnlichen Antworten zu gewährleisten.

## 4. Ergebnisse

### 4.1 Analyse auf Item-Ebene

Zunächst ist eine Betrachtung der Lösungshäufigkeiten der einzelnen Items, gegliedert nach Gruppe mit unterschiedlicher Expertise von Interesse. Diese können Abbildung 2 entnommen werden. Bei sieben Items, also bei etwa der Hälfte der 15 Items, finden wir ein Muster, das grundsätzlich unserer Erwartung entspricht,

dass nämlich mit fortschreitender Expertise auch höhere Lösungshäufigkeiten vorliegen. Die sieben Items sind die zeitliche Verortung des Videoclips in die gesamte Unterrichtsstunde (V2), die Nennung der Handlungsmaßnahmen (V3\_1, V3\_2, V3\_3, V3\_4) sowie die Informationsentnahme zur Weiterführung der Stunde (V4\_2, V4\_3). Eine relative Stärke der Referendare zeichnet sich interessanterweise auf einem Item zur Rechtfertigung einer Handlung ab (V1\_3). Ferner liegen Items vor, bei denen sich kaum Unterschiede zwischen den drei Gruppen zeigen (z. B. V1\_1), aber auch Items, deren Kriterien bei keiner der Gruppen substantiell auftreten, zum Beispiel die komplexeren Anforderungen V1\_4 und V4\_5.

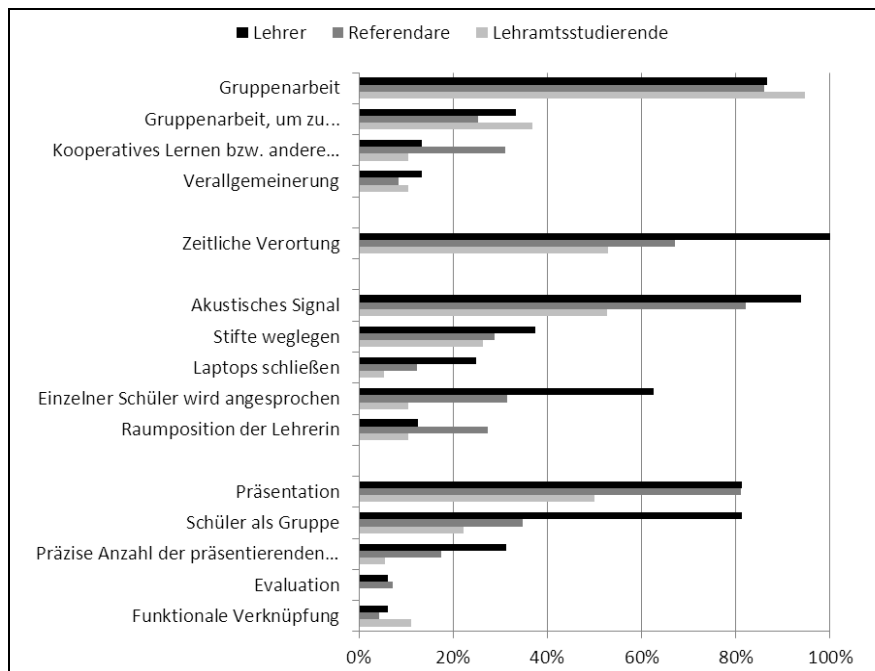


Abbildung 2: Lösungshäufigkeiten der Items in Prozent, gegliedert nach Expertise-Gruppe

#### 4.2 Skalierung der Dimension *Genauigkeit der Wahrnehmung*

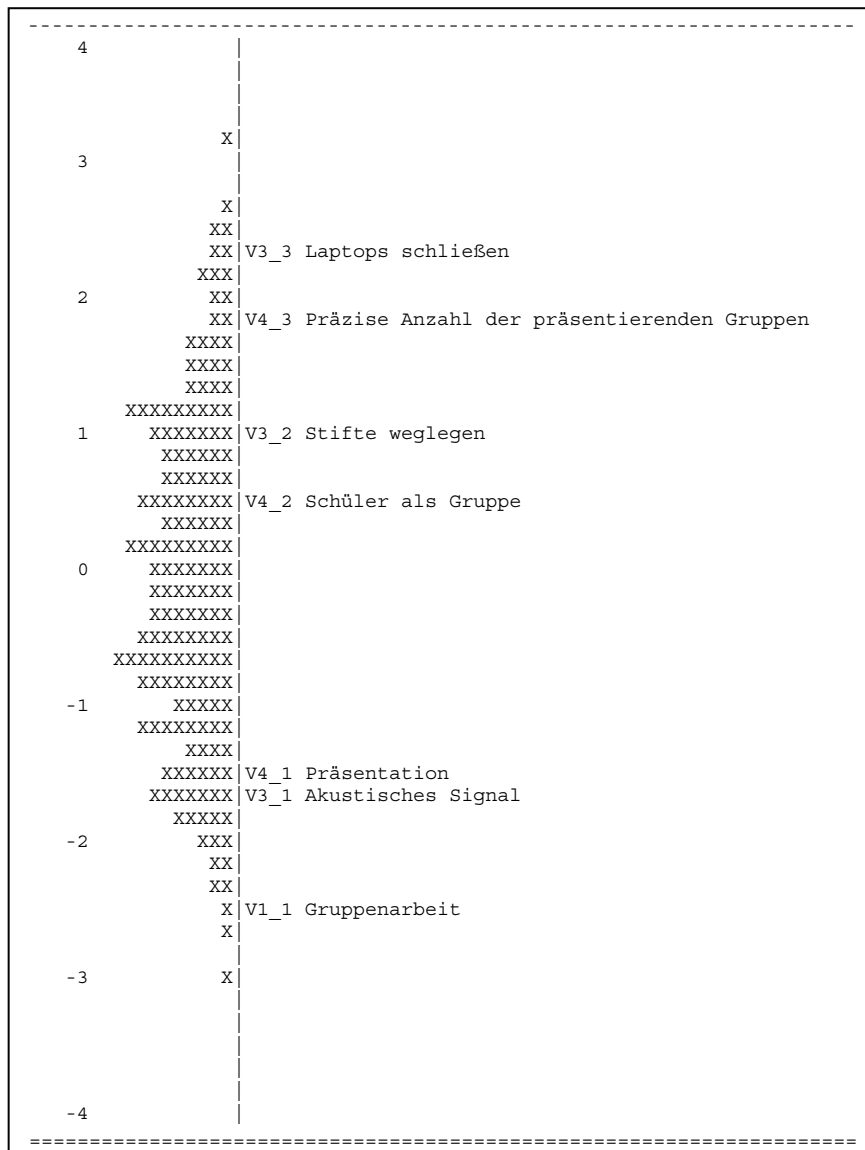
Aus der verwendeten Vignette lassen sich 11 Items beziehen, die über die Genauigkeit der Wahrnehmung durch die Probanden informieren (Tab. 2). Sieben von diesen 11 Items erweisen sich in Skalierungsanalysen als trennscharf und können somit als Indikatoren genutzt werden, um die kognitive Dimension *Genauigkeit*

der Wahrnehmung über eine Skala zu messen. Um Fähigkeitswerte und Item-Schwierigkeitsparameter auf einem Kontinuum abtragen zu können, wurden die sieben Items einer IRT-Skalierung im eindimensionalen Rasch-Modell unterzogen (vgl. Rost, 1996). Verwendet wurde die Software *Conquest* (Wu, Adams & Wilson, 1997). Ergebnisse der Skalierung sind Abbildung 3 und Tabelle 3 zu entnehmen.

Die Reliabilität ist angesichts der geringen Anzahl dichotomer Items, die zudem aus drei unterschiedlichen offenen Testfragen bezogen wurden, akzeptabel (EAP-Reliabilität: 0.60; vergleichbar mit Cronbachs  $\alpha$ ), die Varianz der latenten Variable ist mit 1.63 hingegen als gut zu bezeichnen. Die gewichteten Abweichungsquadrate (*Mean Square*) der Items liegen überwiegend nahe dem Erwartungswert von 1.0 (vgl. Tab. 3) und somit innerhalb des üblicherweise akzeptierten Bereichs (zwischen 0.8 und 1.2; vgl. Wright, Linacre, Gustafsson & Martin-Loff, 1994). Zudem korrelieren die Items jeweils hinreichend mit der Skala (*Discrimination Index* jeweils größer 0.4; vgl. Tab. 3).

Tabelle 3: Item-Schwierigkeitsparameter der Skalierung für die kognitive Anforderungsdimension *Genauigkeit der Wahrnehmung* sowie Angaben zum *Weighted Item Fit* (*Mean Square*, Konfidenzintervall, t-Wert) und zum *Discrimination Index*

Frage	Variable	Bezeichnung	Item-Schwierigkeits-Parameter (SE)	Mean Square (Konfidenzintervall)	t-Wert	Discrimination Index
1	V1_1	Gruppenarbeit	-2.560 (0.214)	1.09 (0.60, 1.40)	0.5	0.41
3	V3_1	Akustisches Signal	-1.769 (0.194)	1.07 (0.72, 1.28)	0.5	0.45
	V3_2	Stifte weglegen	1.060 (0.184)	1.06 (0.78, 1.22)	0.5	0.52
	V3_3	Laptops schließen	2.355 (0.212)	1.02 (0.62, 1.38)	0.1	0.40
4	V4_1	Präsentation	-1.535 (0.192)	0.96 (0.74, 1.26)	-0.3	0.62
	V4_2	Schüler als Gruppe	0.496 (0.180)	0.88 (0.81, 1.19)	-1.3	0.76
	V4_3	Präzise Anzahl der präsentierenden Gruppen	1.954 (0.481)	1.00 (0.68, 1.32)	0.1	0.57



Ein ‚X‘ steht für 0.6 Fälle, die Beschriftungen rechts zeigen die Lage der Items auf dem Kontinuum an

Abbildung 3: Skalierung der Items zur Genauigkeit der Wahrnehmung

Das Item mit der geringsten Schwierigkeit misst die Wahrnehmung von Gruppenarbeit als Funktion der Sitzordnung (Frage 1, Variable V1\_1). Aus der dritten Frage (s. Abb. 1; Tab. 2) resultieren mehrere Items, von denen jedoch nur drei Items in die Skalierung einbezogen werden konnten. Diese betreffen interessanterweise die *unmittelbaren*, und auf die *gesamte* Lerngruppe bezogenen Handlungsmaßnahmen der Lehrerin: das Erteilen eines akustischen Signals (Variable V3\_1) sowie die beiden Anweisungen an alle Schülerinnen und Schüler, die Stifte wegzulegen (Variable V3\_2) und die Laptops zu schließen (Variable V3\_3). Wie den Item-Schwierigkeitsparametern in Tabelle 3 entnommen werden kann, steigen diese für die drei in die Skalierung einbezogenen Items der Frage 3 jeweils um mindestens ein Logit an, was auf eine gute Differenzierung hinsichtlich ihrer Schwierigkeiten verweist. Ähnliches ist bei der Genauigkeit der Wahrnehmung hinsichtlich der Vorhersage der weiteren Handlung festzustellen, welche mit der vierten Frage erfasst werden sollte: Dass eine Präsentation erfolgen soll (Variable V4\_1), ist eine Information, die relativ einfach wahrgenommen werden kann; dass die Präsentation von Schülergruppen geleistet werden soll (Variable V4\_2), erweist sich hingegen als schwieriger wahrzunehmen; und dass es lediglich zwei Gruppen sein werden, die präsentieren sollen bzw. angesichts mangelnder Zeit können (Variable V4\_3), ist eine Information, die einer wesentlich genaueren Wahrnehmung bedarf. Bei den drei entsprechenden Items liegen die Item-Schwierigkeitsparameter rund anderthalb Logits auseinander, was auf eine besonders gute Differenzierung verweist.

#### 4.3 Deskriptive Befunde zur Dimension *Genauigkeit der Wahrnehmung*

Die drei im vorliegenden Beitrag fokussierten Gruppen – Lehramtsstudierende, Referendare, Lehrer – unterscheiden sich in ihren Fähigkeitswerten auf der Skala zur Erfassung der Genauigkeit der Wahrnehmung (s. Abb. 4). Die mittleren Fähigkeitswerte (Median in Abb. 4, arithmetisches Mittel in Tab. 4) steigen erwartungskonform je nach Grad ihrer Expertise von Gruppe zu Gruppe an. Lehramtsstudierende weisen die niedrigsten Ausprägungen auf, Lehrer die höchsten, im Mittelfeld befinden sich die Referendare. In einer univariaten Varianzanalyse mit den Personenfähigkeitswerten als abhängige und der Gruppenzugehörigkeit als unabhängige Variable lassen sich rund elf Prozent der Testleistung durch die Gruppenzugehörigkeit erklären ( $F(2, 105) = 6.66, p < 0.01, \eta^2 = 0.113$ ).<sup>5</sup>

Ordnet man die Mediane bzw. die arithmetischen Mittelwerte der drei Gruppen den Item-Schwierigkeiten der sieben Items in Abbildung 3 zu, so lassen sich die

<sup>5</sup> Dieser Effekt bleibt auch dann statistisch signifikant und von mittlerer praktischer Bedeutsamkeit, wenn das Alter der Probanden kontrolliert wird.

gruppenbezogenen Leistungen wie folgt kriterial interpretieren: Lehramtsstudierende sind in der Regel in der Lage, die Items zur Erfassung einer *einfachen Wahrnehmung* korrekt zu beantworten, d. h. die ersten drei Items V1\_1, V3\_1 und V4\_1. Berufstätige Lehrkräfte hingegen beantworten in der Regel auch Items, welche eine *genauere Wahrnehmung* messen. Referendare können diese schwierigeren Items bereits teilweise korrekt lösen.

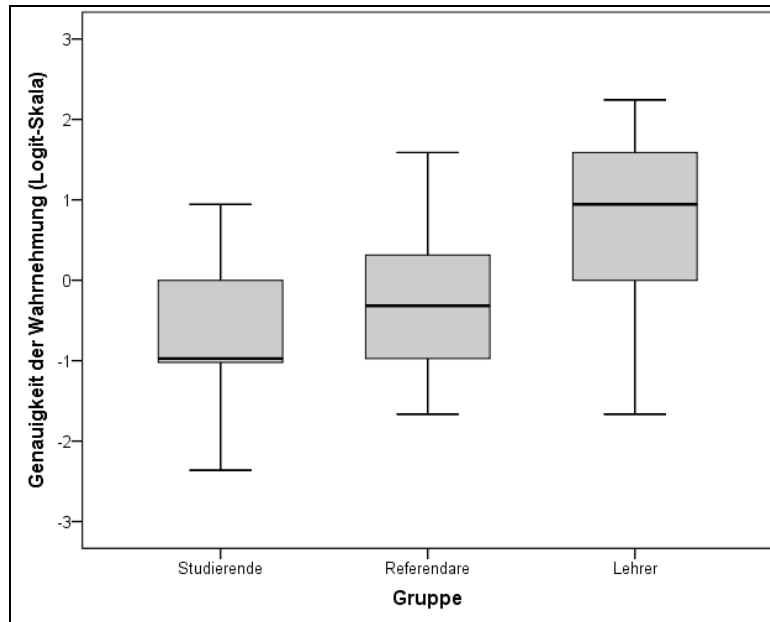


Abbildung 4: Boxplots der Personenfähigkeitswerte zur Dimension *Genauigkeit der Wahrnehmung*, gegliedert nach Expertise-Gruppe

Tabelle 4: Kennwerte zur Dimension *Genauigkeit der Wahrnehmung*, gegliedert nach Expertise-Gruppe

	n	M	SE	SD	Median
Lehramtsstudierende	19	-0.62	0.22	0.96	-0.97
Referendare	73	-0.07	0.10	0.88	-0.32
Lehrer	16	0.54	0.30	1.18	0.94



## 5. Diskussion

Klassenführung (*Classroom Management*) gilt übereinstimmend als zentrale Basisdimension der Unterrichtsqualität (z. B. Helmke, 2009; Wang, Haertel & Walberg, 1993) und das Führen einer Schulklasse stellt an Lehrkräfte hohe berufliche Anforderungen, für deren Bewältigung *Expertise* notwendig ist, sodass für eine erfolgreiche Klassenführung ein fächerübergreifendes, pädagogisches Wissen und Können vorausgesetzt wird (vgl. Punkt 1.1). Im vorliegenden Beitrag wurden Überlegungen zur Messung der *Classroom Management Expertise* (CME) von Lehrkräften dargestellt, die für die Verwendung von Videovignetten sprechen – auch als notwendige Ergänzung eines rein auf dem Papier-Bleistift-Verfahren beruhenden methodischen Zugangs (vgl. König, 2010). In theoretischer Hinsicht wurde die auf die Anforderung der Klassenführung bezogene, wissensgesteuerte Verarbeitung von Unterricht anhand von drei Dimensionen kognitiver Anforderungen konkretisiert: *Genauigkeit der Wahrnehmung*, *Holistische Wahrnehmung* und *Rechtfertigung einer Handlung*. Am Beispiel einer Videovignette und prototypischer Testaufgaben, die diese drei kognitiven Anforderungsdimensionen abbilden, sowie unter Verwendung einer Stichprobe von Lehramtsstudierenden, Referendaren und berufstätigen Lehrpersonen, die Experten-Novizen-Vergleiche ermöglicht, wurde CME anschließend empirisch untersucht.

Die Experten (d. h. berufstätige Lehrkräfte mit durchschnittlich 17 Jahren Berufserfahrung) konnten etwa die Hälfte der Items besser lösen als Referendare, und letztere schnitten wiederum besser ab als Lehramtsstudierende. Dieses Muster ließ sich über eine Skala zur Messung der ersten Dimension (*Genauigkeit der Wahrnehmung*), bestehend aus sieben Items, zusammenfassend und hinreichend reliabel abbilden: Die drei Gruppen unterscheiden sich auf dieser Skala statistisch signifikant mit mittlerer praktischer Bedeutsamkeit. Lehramtsstudierende waren in der Regel in der Lage, die Items zur Erfassung einer *einfachen Wahrnehmung* korrekt zu beantworten, berufstätige Lehrkräfte hingegen beantworteten in der Regel auch Items, welche eine *genauere Wahrnehmung* messen. Referendare konnten bereits teilweise diese schwierigeren Items korrekt lösen. Doch auch ein Item zur Messung der zweiten Dimension (*Holistische Wahrnehmung*) verdeutlichte erwartungskonform, dass berufstätige Lehrkräfte besser als Referendare oder Lehramtsstudierende in der Lage waren, die in der Videovignette gezeigte Unterrichtssituation in das gesamthafte zeitliche Unterrichtsgeschehen einzuordnen.

Eine relative Stärke der Referendare zeichnete sich interessanterweise auf einem Item ab, das wir zur Messung der dritten Dimension (*Rechtfertigung einer Handlung*) einsetzten. Die berufstätigen Lehrkräfte scheinen hinsichtlich dieser Dimension nicht zwangsläufig die elaborierteren Antworten zu geben. Als mögli-

che Ursache könnten die spezifischen mentalen Modelle von Referendaren und berufstätigen Lehrkräften angeführt werden, denen unterschiedlich ausgeprägte Vorwissensstrukturen zugrunde liegen. Im Gegensatz zu berufstätigen Lehrkräften werden Referendare innerhalb ihrer praktischen Ausbildung mit formalem Wissen zur Unterrichtsgestaltung regelmäßig konfrontiert und zur intensiven Einübung, Erklärung und Rechtfertigung sowie zur fortlaufenden Reflexion ihres Handelns unter Prüfungsbedingungen aufgefordert (vgl. König & Blömeke, im Druck). Bei Referendaren ist das vergleichsweise neuartige Wissen noch explizit verbalisierbar und folglich bei salienten Schlüsselreizen abrufbereit. Wie Renkl (1996) ausführt, ist die Verbalisierungsleistung jedoch kein hinreichender Indikator des Handlungswissens. Demzufolge wären die Antworten berufserfahrener Lehrkräfte weniger als das Ergebnis ‚trägen Wissens‘ sondern als ‚inert knowledge‘ zu verstehen: „Knowledge, although seemingly available, is often not used for problem solving. That means it remains ‚inert‘“ (Renkl, Mandl & Gruber, 1996, S. 115). Das spezifische, auf die Anforderungen der Klassenführung hin bezogene Wissen könnte bei erfahrenen Lehrkräften bereits in hohem Maße internalisiert und somit als automatisiert ablaufende Handlungsschemata auf impliziter Ebene konsolidiert sein. Der Transfer von den Wissensinhalten hin zum situationsspezifischen Handeln wäre bereits vollzogen, wodurch sich die Antworten auf Fragebogenitems einem sprachlichen Repräsentationsmodus entziehen (*Expertise Reversal Effect*; vgl. Kalyuga, Ayres, Chandler & Sweller, 2003). So ist zu vermuten, dass eine Messung von CME durchaus auch über geschlossene Antwortformate gelingen kann, wie das verwendete *Multiple-Choice*-Item zur Messung der *Holistischen Wahrnehmung* exemplarisch verdeutlicht. Eine empirische Überprüfung des Zusammenhangs zwischen den Antworten zur Rechtfertigung einer Handlung, der Lehrerfahrung und dem Grad der Automatisierung situationsspezifischen Wissens steht noch aus, wobei insbesondere die Operationalisierung des Automatisierungsgrades weiterführend in den Blick zu nehmen wäre.

Insgesamt sehen wir die hier dargelegte Untersuchung zur Messung von CME in theoretischer Hinsicht anschlussfähig an bisherige Ansätze zur videobasierten Messung von Elementen professioneller Kompetenz (vgl. Punkt 1.1). Unsere Konzeption einer Differenzierung kognitiver Anforderungsdimensionen mit Bezug zur spezifischen beruflichen Aufgabe der Klassenführung erscheint vor dem Hintergrund der hier erzielten empirischen Befunde und Einordnung in den Stand zur Forschung von Lehrerexpertise ein lohnenswerter Ansatz, den wir daher weiter ausarbeiten werden. Einschränkend muss berücksichtigt werden, dass der hier vorgestellte Ansatz zunächst heuristischen Charakter trägt, bedingt durch Art und Umfang der Stichprobe sowie die geringe Anzahl von Items zur Abbildung der zweiten und dritten Dimension. Aus methodischer Perspektive wäre es zum Bei-

spiel sehr wichtig, weitere Videovignetten sowie Variationen zu den hier dargestellten prototypischen Testitems zu entwickeln und eine mehrdimensionale Skalierung anzustreben, sodass reliable Aussagen zu allen drei kognitiven Dimensionen wie auch zu ihren Interkorrelation auf latenter Ebene getroffen werden können. Dies setzt entsprechend große Stichproben voraus, die für die hier dargestellte Pilotstudie aus ressourcenbezogenen Gründen nicht realisiert werden konnten (z. B. kein Einsatz von *Incentives* zur Rekrutierung von Probanden). Zu konstatieren bleibt also, dass die hier berichteten empirischen Befunde in ihrer Aussagekraft entsprechend begrenzt sind und z. B. Verallgemeinerungen nur bedingt möglich sind.

Mit unserer hier skizzierten Forschungsperspektive ist allerdings die Chance verbunden, handlungsbezogene Kognitionen von Lehrkräften in verstärktem Maße in den Blick zu nehmen als es bislang mit Papier-Bleistift-Verfahren möglich war, sodass ein erweiterter Zugang zur Abbildung und Modellierung von CME eröffnet wird. Insbesondere unter Aspekten der Konstruktvalidität wäre die Bestimmung des Zusammenhangs der hier gemessenen CME zu bisherigen Maßen der Kompetenzmessung wie pädagogisches Unterrichtswissen (König, 2012) sehr aufschlussreich. Ein weiteres zentrales Forschungsanliegen besteht darin, zu untersuchen, welche Bedeutung die so gemessene CME als Element professioneller Kompetenz von Lehrkräften für die Qualität des von ihnen gehaltenen Unterrichts besitzt. Eine entsprechende Studie, die damit auch wichtigen Einblick in die Validität des Konstrukts geben kann, wird von uns derzeit vorbereitet, sodass in absehbarer Zeit Erkenntnisse zum Zusammenhang zwischen CME von Lehrkräften und der Qualität ihres Unterrichts vorliegen werden.

### Literatur

- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft, 9, 469-520.
- Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV (S. 29-53). Münster: Waxmann.
- Blömeke, S., Bremerich-Vos, A., Haudeck, H., Kaiser, G., Lehmann, R., Nold, G., Schwippert, K. & Willenberg, H. (Hrsg.). (2011). Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen. Erste Ergebnisse aus TEDS-LT. Münster: Waxmann.
- Blömeke, S., Kaiser, G. & König, J. (2009). Längsschnittliche Entwicklung der Kompetenzen von Junglehrkräften: Follow-Up zur internationalen Vergleichsstudie TEDS-M (TEDS-FU). Antrag auf Gewährung einer Sachbeihilfe an die Deutsche Forschungsgemeinschaft.

- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2010a). TEDS-M 2008 – Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich. Münster: Waxmann.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2010b). TEDS-M 2008 – Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich. Münster: Waxmann.
- Bromme, R. (1992). Der Lehrer als Experte: zur Psychologie des professionellen Wissens. Bern: Huber.
- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In Weinert, F. E. (Hrsg.), Enzyklopädie der Psychologie: Psychologie des Unterrichts und der Schule (Bd. 3) (S. 177-212). Göttingen.
- Bromme, R. (2001). Teacher expertise. *International Encyclopedia of the social & behavioral sciences* (pp. 15459-15465). Amsterdam: Smelser and Baltes.
- Carter, K., Sabers, D. S., Cushing, K. S., Pinnegar, P. & Berliner, D. C. (1987). Processing and using information about students: A study of expert, novice and postulant teachers. *Teaching and Teacher Education*, 3, 147-157.
- Copeland, W. D. (1987). Classroom management and student teachers' cognitive abilities: A relationship. *American Educational Research Journal*, 24, 219-236.
- Doyle, W. (1985). Recent research on classroom management: Implications for teacher preparation. *Journal of Teacher Education*, 36, 31-35.
- Doyle, W. (2006). Ecological approaches to classroom management. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 97-125). Mahwah, NJ: Erlbaum.
- Es, E. A. van & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education*, 24, 244-276.
- Evertson, C. M. & Weinstein, C. S. (Eds.). (2006). *Handbook of classroom management: Research, practice, and contemporary Research*. Mahwah, NJ: Erlbaum.
- Fennema, E. & Franke, M. L. (1992). Teachers' knowledge and its impact. In D. A. Grouws, (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 147-164). New York: Macmillan Publishing Company.
- Fenstermacher, G. D. (1994). The knower and the known: The nature of knowledge in research on teaching. In L. Darling-Hammond (Ed.), *Review of research in education* (pp. 3-56). Washington: American Educational Research Association.

- Gold, B., Meschede, N., Wolters, M., Möller, K., Holodynski, M., Steffensky, M., Glaser, O. & Lehmann, T. (2011). Kompetenzen zur Analyse der Lernwirksamkeit von naturwissenschaftlichem Grundschulunterricht – Theoretische Modellierung und empirische Erfassung. Poster auf dem Jahrestreffen des BMBF-Forschungsschwerpunktes „Entwicklung von Professionalität des pädagogischen Personals in Bildungseinrichtungen“, 07.-08. Juli 2011, Wuppertal.
- Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. New York: Teachers College Press.
- Gruber, H., Harteis, C. & Rehl, M. (2006). Professional Learning: Erfahrung als Grundlage von Handlungskompetenz. *Bildung und Erziehung*, 59, 193-203.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Kallmeyer.
- Kalyuga, S., Ayres, P., Chandler, P. & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38, 23-31.
- Klieme, E. (2006). Empirische Unterrichtsforschung: aktuelle Entwicklungen, theoretische Grundlagen und fachspezifische Befunde. Einführung in den Thementeil. *Zeitschrift für Pädagogik*, 52, 765-773.
- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. *Zeitschrift für Pädagogik*, 6, 876-903.
- König, J. (2010). Lehrerprofessionalität - Konzepte und Ergebnisse der internationalen und deutschen Forschung am Beispiel fachübergreifender, pädagogischer Kompetenzen. In J. König & B. Hofmann (Hrsg.), *Professionalität von Lehrkräften – Was sollen Lehrkräfte im Lese- und Schreibunterricht wissen und können?* (S. 40-105). Berlin: DGLS.
- König, J. (2012). Die Entwicklung von pädagogischem Unterrichtswissen: Theoretischer Rahmen, Testinstrument, Skalierung und Ergebnisse. In J. König & Seifert, A. (Hrsg.), *Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerbildung* (S.143-182). Münster: Waxmann.
- König, J. & Blömeke, S. (2009). Pädagogisches Wissen von angehenden Lehrkräften: Erfassung und Struktur von Ergebnissen der fachübergreifenden Lehrerbildung. *Zeitschrift für Erziehungswissenschaft*, 12, 499-527.
- König, J. & Blömeke, S. (im Druck). Preparing teachers of mathematics in Germany. In J. Schulle, L. Ingvarson & R. Holdgreve-Resendez (Eds.), *TEDS-M Encyclopaedia. A guide to teacher education context, structure and quality assurance in the seventeen TEDS-M countries*.

- König, J., Blömeke, S., Paine, L., Schmidt, W. H. & Hsieh, F. J. (2011). General pedagogical knowledge of future middle school teachers: On the complex ecology of teacher education in the United States, Germany, and Taiwan. *Journal of Teacher Education*, 62, 188-201.
- König, J. & Seifert, A. (Hrsg.). (2012). *Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerbildung*. Münster: Waxmann.
- Kounin, J. S. (1976). *Techniken der Klassenführung*. Stuttgart: Waxmann.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Leinhardt, G. & Greeno, J. G. (1986). The cognitive skill of teaching. *Journal of Educational Psychology*, 78, 75-95.
- Neuweg, G. H. (2008). Grundlagen der Lehrer/innen/kompetenz. *Educational Sciences (Odgojne znanosti)*, 10, 13-22.
- Neuweg, G. H. (2011). Das Wissen der Wissensvermittler. Problemstellungen, Befunde und Perspektiven der Forschung zum Lehrwissen. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 451-477). Münster: Waxmann.
- Oser, F. & Patry, J.-L. (1994). Sichtstruktur und Basismodelle des Unterrichts: Über den Zusammenhang von Lehren und Lernen unter dem Gesichtspunkt psychologischer Lernverläufe. In R. Olechowski & B. Rollett (Hrsg.), *Theorie und Praxis. Aspekte empirisch-pädagogischer Forschung* (S. 138-146). Bern: Lang.
- Plöger, W. & Scholl, D. (2011). Unterrichts- und Lehrerexpertise – Perspektiven für ihre künftige Erforschung. *Pädagogische Rundschau*, 65, 647-663.
- Putnam, R. T. & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning. *Educational Researcher*, 29, 4-15.
- Renkl, A. (1996). Träges Wissen, Wenn Erlerntes nicht genutzt wird. *Psychologische Rundschau*, 47, 78-92.
- Renkl, A., Mandl, H. & Gruber, H. (1996). Inert knowledge: Analyses and remedies. *Educational Psychologist*, 31, 115-121.
- Rost, J. (1996). *Lehrbuch Testtheorie, Testkonstruktion*. Bern: Huber.
- Sabers, D. S., Cushing, K. S. & Berliner, D. C. (1991). Differences among teachers in a task characterized by simultaneity, multidimensionality, and immediacy. *American Educational Research Journal*, 28, 63-88.

- Schaper, N. (2009). Aufgabenfelder und Perspektiven bei der Kompetenzmodellierung und -messung in der Lehrerbildung. *Lehrerbildung auf dem Prüfstand*, 2, 166-199.
- Seidel, T., Blomberg, G. & Stürmer, K. (2010). „Observer“ – Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. *Zeitschrift für Pädagogik*, 56, Beiheft, 296-306.
- Shavelson, R. J. & Stern, O. (1981). Research on teacher's pedagogical thoughts, judgements, decisions, and behavior. *Review of Educational Research*, 51, 455-498.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Havard Educational Research*, 57, 1-22.
- Voss, T. & Kunter, M. (2011). Pädagogisch-psychologisches Wissen von Lehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 193-214). Münster: Waxmann.
- Wang, M. C., Haertel, G. D. & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63, 249-294.
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and selecting key competencies* (pp. 45-66). Göttingen: Hogrefe.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.
- Wright, B. D., Linacre, M., Gustafsson, J.-E. & Martin-Loff, P. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8, 370.
- Wu, M. L., Adams, R. J. & Wilson, M. R. (1997). *ConQuest: Multi-aspect test software [computer program]*. Camberwell, Vic.: Australian Council for Educational Research.

*Anschrift der Autoren:*

Prof. Dr. Johannes König und Dr. Morena Lebens, Institut für Allgemeine Didaktik und Schulforschung, Humanwissenschaftliche Fakultät der Universität zu Köln, Gronewaldstraße 2, 50931 Köln, E-Mail: johannes.koenig@uni-koeln.de